A NEXT-GENERATION FRAMEWORK FOR AI-AUGMENTED, SUBMISSION-GRADE LITERATURE REVIEWS: A MULTI-PROJECT ANALYSIS OF SPEED, ACCURACY, AND WORKFLOW EVOLUTION

ISPOR 2026

MadeAi

SA38

AUTHOR(S)

Angeline Dhas, BS[1] | **Revanth M,** B.E[2] | **Diwyashri Govindarajaperumal,** B.Pharm[2] | **Siva Karthick Bagavathiappan,** B.E.[3] | **Vince Salerno,** PharmD, RPh[1] | **Viji Queen,** PharmD[2]

[1]MadeAi, Cambridge, MA, USA, [2]MadeAi, Nagercoil, India, [3]MadeAi, Etobicoke, ON, Canada.

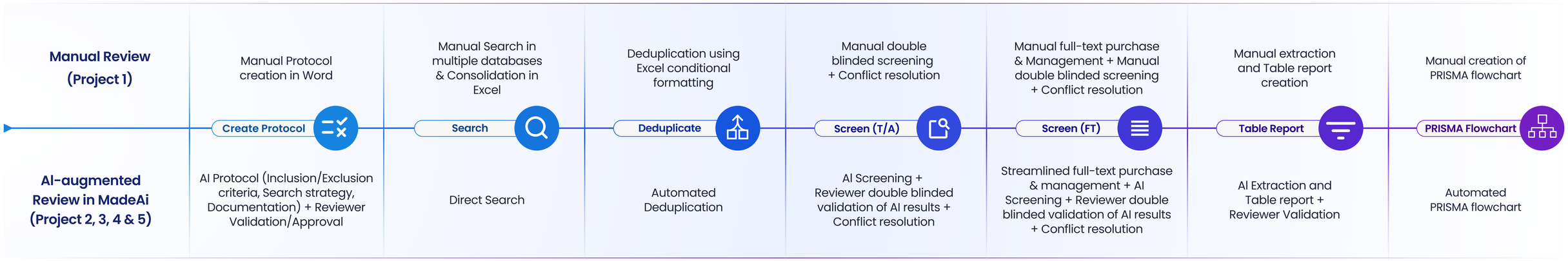
Abstract

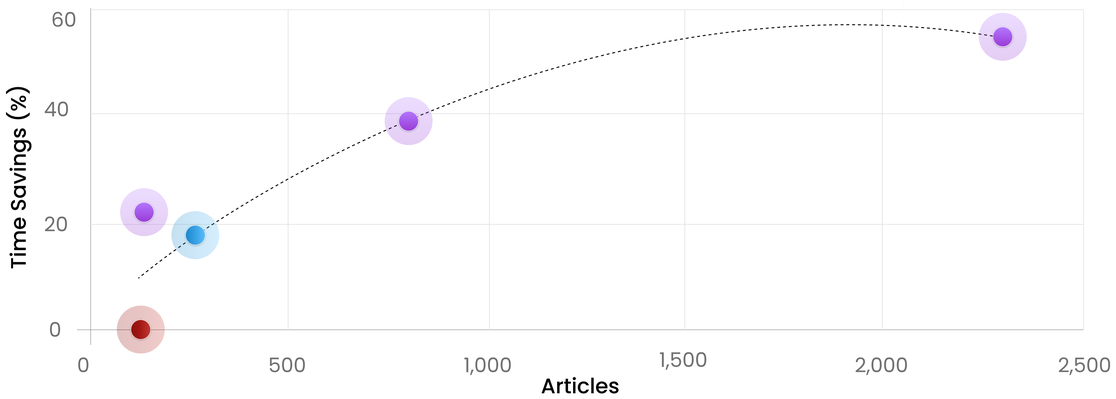
Objectives

Although AI has been applied to literature reviews for several years, adoption within regulatory-oriented evidence workflows remains limited due to transparency, reproducibility, traceability, and methodological concerns. This study evaluates a standardized, submission-grade framework for AI-augmented literature reviews, assessing its impact on speed, accuracy, workflow efficiency, and ROI across multiple real-world regulatory projects.

Methods

A comparative analysis was conducted across five regulatory literature reviews, supporting CERs and PSURs. One project was completed using a fully manual workflow as a baseline, while four projects applied an AI-augmented methodology using a structured framework. All followed an identical process, including protocol development, database searching, deduplication, title/abstract screening, full-text screening, data appraisal, evidence table creation, and PRISMA generation. In AI-augmented projects, AI was incorporated across all stages except the baseline. Human reviewers adjudicated screening decisions, verified extracted data, and performed final quality control. Metrics included time spent at each stage, AI-output accuracy, workflow efficiency, traceability, and ROI.





Project	Total Volume	AI Accuracy	Human Hours	Time Savings
Manual Project	120	NA	71	0%
During AI Adoption	265	88%	130	17%
Post AI Adoption	130	90%	60	22%
	2,300	89%	610	55%
	800	86%	290	39%

Results

The AI-augmented framework demonstrated progressively increasing time savings, achieving a 32% time savings (initial adoption project), which increased to 33-51% in subsequent projects, highlighting sustained gains as workflows matured. Time savings increased even as article volume expanded nearly 19 times (120 to 2,300 articles), demonstrating strong scalability and hours-per-article efficiency. AI reduced the manual effort in screening, data extraction, and evidence table drafting, while human reviewers maintained regulatory accuracy and oversight. AI accuracy across protocol development, screening, and data extraction remained consistently high (84-90%), enabling reliable outputs with reduced manual effort and shorter quality-control cycles. Human reviewer oversight ensured regulatory accuracy and methodological defensibility throughout all projects.

Conclusion

This multi-project evaluation demonstrates that a structured, transparent, AI-augmented framework can deliver substantial efficiency gains while maintaining accuracy and regulatory rigor. The approach enables scalable, defensible adoption of AI in submission-grade literature reviews and supports its integration into regulatory evidence-generation workflows.